

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:33 ; Search time 37.5 seconds  
(without alignments)  
60.181 Million cell updates/sec

Title: US-09-331-631a-24\_COPY\_29\_94  
Perfect score: 382  
Sequence: 1 HDEDDRRGSHSLQCVQRC.....EOEEQGRGRMGEGERE 66

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues  
Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	382	100.0	637	19	W62837
2	138	36.1	593	19	W62835
3	115	30.1	666	19	W62828
4	105	27.5	666	19	W62829
5	93	24.3	590	19	W62832
6	91	23.8	625	19	W62830
7	75	19.6	699	21	Y67598
8	74.5	19.5	1856	20	Y21802
9	74.5	19.5	1856	21	Y83370
10	73	19.1	605	19	W62838
11	73	19.1	605	20	Y40399
12	72	18.8	525	19	W62831

13	72	18.8	566	13	R20181
14	70	18.3	156	20	V76583
15	70	18.3	1122	21	V32218
16	70	18.3	1175	21	V32219
17	70	18.3	1218	21	V32217
18	69	18.1	1162	21	V58500
19	67.5	17.7	1382	18	W31867
20	67	17.5	409	20	W90342
21	67	17.5	489	20	W90341
22	65.5	17.1	1898	20	V30795
23	65	17.0	444	20	W90340
24	65	17.0	524	20	W90339
25	65	17.0	1039	19	W73309
26	64.5	16.9	242	14	R31229
27	64.5	16.9	242	14	W68538
28	63.5	16.6	575	19	W58856
29	63	16.5	147	19	W72404
30	63	16.5	155	19	W72382
31	63	16.2	611	20	V29039
32	62	16.2	492	20	V07742
33	61	16.0	594	14	R34936
34	61	16.0	1031	19	W56102
35	60.5	15.8	1088	21	V69192
36	60	15.7	539	19	W33628
37	60	15.7	584	19	W78916
38	59.5	15.6	157	20	V74172
39	59	15.4	395	17	W03474
40	59	15.4	1132	17	R97866
41	58.5	15.3	306	21	V51614
42	58.5	15.3	516	7	P61362
43	58.5	15.3	1244	18	W19785
44	58	15.2	197	17	W00399
45	58	15.2	203	17	W06107

## ALIGNMENTS

RESULT 1	W62837	standard; Protein; 637 AA.
XX	W62837;	
AC	W62837;	
XX	27-OCT-1998	(first entry)
DT	XX	
DE	Hordeum vulgare antimicrobial protein.	
XX	antimicrobial protein; infestation; control.	
KW	XX	
OS	Hordeum vulgare.	
XX	XX	
PN	W09827805-A1.	
XX	02-JUL-1998.	
PD	XX	
XX	22-DEC-1997;	97WO-AU00874.
PF	XX	
PR	20-DEC-1996;	96AU-0004275.
XX	XX	
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	
PI	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;	
DR	WPI: 1998-377279/32.	
XX	Novel anti-microbial protein from e.g. Macadamia integrifolia -	
PT	useful for controlling microbial infestations of plants or mammals	
XX	Claim 1; Page 60-62; 96pp; English.	
PS	The sequence is that of an antimicrobial protein which can	
CC	be used to control microbial infestations in plants and mammalian	

Sequence encoded b  
Human ovarian tumor  
Drosophila PER B p  
Drosophila PER C p  
Drosophila PER A p  
HHV8 ORF 73 protein  
Human metastasis-a  
G. max truncated S  
G. max SBP2 protein  
A human trichohyal  
G. max truncated S  
G. max SBP1 protein  
Cellulose synthase  
Prepro-thyrotropin  
Pepto TRH protein.  
C. acidivorans gam  
Arabidopsis thaliana  
T. gondii immunog  
Coffee storage pro  
CENP-B. Homo sapi  
Euplotes telomeras  
A human monocyste-m  
Yeast transcription  
Bovine butyrophilin  
Human prostate tumor  
Mouse SRY-related  
Chicken leucocytos  
S. avermitilis HRP  
Soybean glycinin A  
Mouse multiple reg  
Human myelin basic  
Foetal myelin basl

CC animals.  
XX  
SQ Sequence 637 AA;

Query Match 100.0%; Score 382; DB 19; Length 637;  
Best Local Similarity 100.0%; Pred. No. 8.9e-39;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HDDEDDRGGHSLQOCVQRCQERPRYSARCVQECRDDQOQHGRHEQEEQGRGWMH 60  
DB 29 dddeddrirghslqgcqrerqerprysarvcvqecrddqgqghrheqeeqgrgrgwmh 88

OY 61 EGEREE 66  
DB 89 eger ee 94

## RESULT 2

W62835  
ID W62835 standard; Protein; 593 AA.

XX  
AC W62835;

XX  
DF 27-OCT-1998 (first entry)

DE Zea mays antimicrobial protein.

XX  
KM antimicrobial protein; Infestation; control.

OS Zea mays.

XX  
PN W09827805-A1.

XX  
PD 02-JUL-1998.

XX  
PF 22-DEC-1997; 97MO-AU00874.

XX  
PR 20-DEC-1996; 96AU-0004275.

XX  
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

XX  
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;

XX  
DR WPI; 1998-377279/32.

XX  
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -

XX  
PS useful for controlling microbial infestations of plants or mammals

XX  
PS Claim 1; Page 58-60; 96pp; English.

XX  
CC The sequence is that of an antimicrobial protein which can  
CC be used to control microbial infestations in plants and mammalian  
XX animals.

XX  
SQ Sequence 593 AA;

Query Match 36.1%; Score 138; DB 19; Length 593;  
Best Local Similarity 37.7%; Pred. No. 5.2e-09;

Matches 26; Conservative 14; Mismatches 21; Indels 8; Gaps 3;

OY 2 DDEDDRGGHSLQOCVQRCQERPRYSARCVQECRDDQOQHGRHEQEEQGRGWMH 57  
DB 26 dddeddrirghslqgcqrerqerprysarvcvqecrddqgqghrheqeeqgrgrgwmh 82

OY 58 WHEGEREE 66  
DB 83 -sseder ee 90

RESULT 3  
W62828

ID W62828 standard; Protein; 666 AA.

XX  
AC W62828;

XX  
DT 27-OCT-1998 (first entry)

XX  
DE Macadamia integrifolia antimicrobial protein.

XX  
KM antimicrobial protein; Infestation; control.

XX  
OS Macadamia integrifolia.

XX  
FH Key Location/Qualifiers

FT Peptide 1..28 /note= "signal peptide"

FT Protein 29..666 /note= "mature protein"

XX  
PN W09827805-A1.

XX  
PD 02-JUL-1998.

XX  
PF 22-DEC-1997; 97MO-AU00874.

XX  
PR 20-DEC-1996; 96AU-0004275.

XX  
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

XX  
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;

XX  
DR WPI; 1998-377279/32.

XX  
DR N-PSDB; V42310.

XX  
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -

XX  
PT useful for controlling microbial infestations of plants or mammals

XX  
PS Claim 1; Page 34-36; 96pp; English.

XX  
CC The sequence is that of an antimicrobial protein which can

XX  
CC be used to control microbial infestations in plants and mammalian

XX  
CC animals.

XX  
SQ Sequence 666 AA;

Query Match 30.1%; Score 115; DB 19; Length 666;  
Best Local Similarity 37.3%; Pred. No. 3.9e-06;

Matches 25; Conservative 12; Mismatches 26; Indels 4; Gaps 3;

OY 3 DDEDDRGGHSLQOCVQRCQERPRYSARCVQECRDDQOQHGRHEQEEQGRGWMH 59  
DB 182 eednkdprgqreyedrrrceqgeqgrqh-qcqlrcreqgrgqgdmnpqrgsgsgr 240

OY 60 GEGEREE 66  
DB 241 eeg eeeq 247

## RESULT 4

W62829  
ID W62829 standard; Protein; 666 AA.

XX  
AC W62829;

XX  
DT 27-OCT-1998 (first entry)

XX  
DE Macadamia integrifolia antimicrobial protein.

XX  
KM antimicrobial protein; Infestation; control.

XX  
OS Macadamia integrifolia.

XX  
FH Key Location/Qualifiers

FT	Peptide	1..28
FT		/note= "signal peptide"
FT	Protein	29..666
FT		/note= "mature protein"
XX		
PN	MO9827805-A1.	
XX		
PD	02-JUL-1998.	
XX		
PE	22-DEC-1997;	97MO-AU00874.
XX		
PR	20-DEC-1996;	96AU-0004275.
XX		
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	
XX		
PI	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;	
XX		
DR	WPI: 1998-377279/32.	
DR	N-PSDB: V42311.	
XX		
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia -	
PT	useful for controlling microbial infestations of plants or mammals	
XX		
PS	Claim 1; Page 39-41; 96pp; English.	
XX		
CC	The sequence is that of an antimicrobial protein which can	
CC	be used to control microbial infestations in plants and mammalian	
CC	animals.	
XX		
SO	Sequence 666 AA;	

```

Query Match 27.5%; Score 105; DA 19; Length 666;
Best Local Similarity 34.3%; Pred. No. 6,4e-05;
Matches 23; Conservative 15; Mismatches 25; Indels 4; Gaps

QY 3 DEDDRRG-GHSLDQCVCORCHQERPRYSHARCVCQECRDDQOQNGR-HEQEEDGGRGWH 59
Db 182 eedhkrdpqgqeyedcrrrccqgqeprrgy-qcqrtrcreqrqhqrgdlnprqgsgrgy 240
QY 60 GEGERE 66
Db 241 eegeekq 247

RESULT 5
ID W62832 standard; Protein; 590 AA.
XX W62832;
XX AC W62832;
XX DT 27-OCT-1998 (first entry)
XX DE Gossypium hirsutum antimicrobial protein.
XX KW antimicrobial protein; infestation; control.
XX OS Gossypium hirsutum.
XX PN W09827805-A1.
XX PD 02-JUL-1998.
XX PF 22-DEC-1997; 97WO-AU00874.
XX PR 20-DEC-1996; 96AU-0004275.
XX PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX DR WPI: 1998-377279/32.
XX

```

PT Novel anti-microbial protein from e.g. *Macadamia integrifolia* -  
 PT useful for controlling microbial infestations of plants or mammals  
 XX  
 XX  
 PS Claim 1; Page 49-51; 96pp; English.  
 XX  
 CC The sequence is that of an antimicrobial protein which can  
 CC be used to control microbial infestations in plants and mammalian  
 CC animals.  
 XX  
 XX Sequence 590 AA;  
 SO

QY	3	DEDDRRGSHLQOCVORCRQRP-R-YSHARCVOECRDQQOH--GRHEOEFGRGKGW	58
Dd	114	egeqgsgfgrfgceqchqhgcqgqrpekkxqcrcrckkygenpwrgyeteaeete-	171
QY	59	HGEGRRE	66
Dd	172	--egeqeg	177
RESULT	6		
W62830			
ID	W62830	standard; Protein; 625 AA.	
AC	W62830;		
XX	27-OCT-1998	(first entry)	
DT			
XX	Macadamia integrifolia antimicrobial protein.		
De			
XX	antimicrobial protein; infestation; control.		
KM			
XX	Macadamia integrifolia.		
OS			
XX	Key	Location/Qualifiers	
FH	Peptide	1..28	
FT		/note= "signal peptide"	
FT	Protein	29..666	
FT		/note= "mature protein"	
XX			
PN	WC9827805-A1.		
PD	02-JUL-1998.		
XX			
Pf	22-DEC-1997;	97WO-AU00874.	
XX			
PR	20-DEC-1996;	96AU-0004275.	
XX			
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.		
PI	Bower NI, Goulter KC, Green JT, Manners JM, Marcus JP;		
XX			
DR	WPL: 1998-377279/32.		
N-PSDB:	V42316.		
XX			
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia -		
PT	useful for controlling microbial infestations of plants or mammals		
XX			
PS	Claim 1; Page 43-45; 96pp; English.		
CC	The sequence is that of an antimicrobial protein which can		
CC	be used to control microbial infestations in plants and mammalian		
CC	animals.		
SO	Sequence	625 AA;	
Query Match	23.8%	Score 91; DB 19; Length 625;	



XX Bacillus subtilis.  
OS  
XX Key Location/Qualifiers  
FT Misc-difference 67 /note= "Encoded by TAA stop codon"  
FT Misc-difference 91 /note= "Encoded by TAA stop codon"  
FT Misc-difference 92 /note= "Encoded by TAA stop codon"  
FT Misc-difference 111 /note= "Encoded by TAA stop codon"  
FT Misc-difference 120 /note= "Encoded by TGA stop codon"  
FT Misc-difference 133 /note= "Encoded by TAG stop codon"  
FT Misc-difference 146 /note= "Encoded by TAG stop codon"  
FT Misc-difference 151 /note= "Encoded by TAA stop codon"  
FT Misc-difference 176 /note= "Encoded by TGA stop codon"  
FT Misc-difference 184 /note= "Encoded by TGA stop codon"  
FT Misc-difference 201 /note= "Encoded by TAG stop codon"  
FT Misc-difference 208 /note= "Encoded by TAA stop codon"  
FT Misc-difference 234 /note= "Encoded by TAA stop codon"  
FT Misc-difference 245 /note= "Encoded by TGA stop codon"  
FT Misc-difference 253 /note= "Encoded by TGA stop codon"  
FT Misc-difference 269 /label= Encoded by TAA stop codon  
FT Misc-difference 295 /note= "Encoded by TAA stop codon"  
FT Misc-difference 298 /note= "Encoded by TAG stop codon"  
FT Misc-difference 323 /note= "Encoded by TGA stop codon"  
FT Misc-difference 324 /note= "Encoded by TGA stop codon"  
FT Misc-difference 331 /note= "Encoded by TGA stop codon"  
FT Misc-difference 341 /note= "Encoded by TAA stop codon"  
FT Misc-difference 344 /note= "Encoded by TAA stop codon"  
FT Misc-difference 356 /note= "Encoded by TAA stop codon"  
FT Misc-difference 358 /note= "Encoded by TGA stop codon"  
FT Misc-difference 413 /note= "Encoded by TAA stop codon"  
FT Misc-difference 469 /note= "Encoded by TGA stop codon"  
FT Misc-difference 486 /note= "Encoded by TAA stop codon"  
FT Misc-difference 501 /note= "Encoded by TGA stop codon"  
FT Misc-difference 523 /note= "Encoded by TGA stop codon"  
FT Misc-difference 532 /note= "Encoded by TGA stop codon"  
FT Misc-difference 586 /note= "Encoded by TGA stop codon"  
FT Misc-difference 593 /note= "Encoded by TGA stop codon"  
FT Misc-difference 596 /note= "Encoded by TGA stop codon"  
FT Misc-difference 632 /note= "Encoded by TAA stop codon"

FT Misc-difference 640 /note= "Encoded by TGA stop codon"  
FT Misc-difference 681 /note= "Encoded by TGA stop codon"  
FT Misc-difference 695 /note= "Encoded by TAA stop codon"  
FT Misc-difference 710 /note= "Encoded by TGA stop codon"  
FT Misc-difference 717 /note= "Encoded by TGA stop codon"  
FT Misc-difference 731 /note= "Encoded by TGA stop codon"  
FT Misc-difference 747 /note= "Encoded by TGA stop codon"  
FT Misc-difference 756 /note= "Encoded by TGA stop codon"  
FT Misc-difference 764 /note= "Encoded by TAA stop codon"  
FT Misc-difference 775 /note= "Encoded by TAG stop codon"  
FT Misc-difference 780 /note= "Encoded by TGA stop codon"  
FT Misc-difference 804 /note= "Encoded by TGA stop codon"  
FT Misc-difference 807 /note= "Encoded by TAA stop codon"  
FT Misc-difference 813 /note= "Encoded by TAA stop codon"  
FT Misc-difference 834 /note= "Encoded by TAA stop codon"  
FT Misc-difference 854 /note= "Encoded by TAA stop codon"  
FT Misc-difference 860 /note= "Encoded by TAA stop codon"  
FT Misc-difference 866 /note= "Encoded by TAA stop codon"  
FT Misc-difference 875 /note= "Encoded by TAA stop codon"  
FT Misc-difference 877 /note= "Encoded by TGA stop codon"  
FT Misc-difference 882 /note= "Encoded by TAA stop codon"  
FT Misc-difference 887 /note= "Encoded by TGA stop codon"  
FT Misc-difference 891 /note= "Encoded by TGA stop codon"  
FT Misc-difference 910 /note= "Encoded by TAA stop codon"  
FT Misc-difference 916 /note= "Encoded by TGA stop codon"  
FT Misc-difference 942 /note= "Encoded by TGA stop codon"  
FT Misc-difference 975 /note= "Encoded by TAA stop codon"  
FT Misc-difference 976 /note= "Encoded by TGA stop codon"  
FT Misc-difference 981 /note= "Encoded by TGA stop codon"  
FT Misc-difference 997 /note= "Encoded by TGA stop codon"  
FT Misc-difference 1013 /note= "Encoded by TAA stop codon"  
FT Misc-difference 1020 /note= "Encoded by TGA stop codon"  
FT Misc-difference 1026 /note= "Encoded by TGA stop codon"  
FT Misc-difference 1028 /note= "Encoded by TGA stop codon"  
FT Misc-difference 1044 /note= "Encoded by TAA stop codon"  
FT Misc-difference 1070 /note= "Encoded by TGA stop codon"  
FT Misc-difference 1099 /note= "Encoded by TGA stop codon"  
FT /note= "Encoded by TGA stop codon"

```
FT Misc-difference 1102 /note- "Encoded by TGA stop codon"
FT Misc-difference 1114 /note- "Encoded by TGA stop codon"
FT Misc-difference 1118 /note- "Encoded by TGA stop codon"
FT Misc-difference 1127 /note- "Encoded by TGA stop codon"
FT Misc-difference 1130 /note- "Encoded by TGA stop codon"
FT Misc-difference 1143 /note- "Encoded by TGA stop codon"
FT Misc-difference 1159 /note- "Encoded by TAA stop codon"
FT Misc-difference 1161 /note- "Encoded by TGA stop codon"
FT Misc-difference 1167 /note- "Encoded by TGA stop codon"
FT Misc-difference 1172 /note- "Encoded by TGA stop codon"
FT Misc-difference 1179 /note- "Encoded by TAA stop codon"
FT Misc-difference 1212 /note- "Encoded by TGA stop codon"
FT Misc-difference 1230 /note- "Encoded by TGA stop codon"
FT Misc-difference 1255 /note- "Encoded by TGA stop codon"
FT Misc-difference 1260 /note- "Encoded by TAA stop codon"
FT Misc-difference 1267 /note- "Encoded by TAA stop codon"
FT Misc-difference 1273 /note- "Encoded by TGA stop codon"
FT Misc-difference 1305 /note- "Encoded by TGA stop codon"
FT Misc-difference 1313 /note- "Encoded by TAA stop codon"
FT Misc-difference 1334 /note- "Encoded by TGA stop codon"
FT Misc-difference 1334 /note- "Encoded by TAA stop codon"
FT Misc-difference 1517 /note- "Encoded by TAA stop codon"
FT Misc-difference 1529 /note- "Encoded by TAA stop codon"

Query Match 19.5%; Score 74.5; DB 21; Length 1856;
Best Local Similarity 22.8%; Pred. No. 1.1;
Matches 21; Conservative 10; Mismatches 34; Indels 27; Gaps 3;

QY 2 DDDDDR-----RGHSIQOCYQRCRQERPRYSNARC-VQECRDQOQHGRHEQ 49
DB 477 dqhderwhxkggrhpggrpgeaeekksalayediprvrhaksgcpxrtdsgyghxqgmd 536
QY 50 EEOG-----RGRGMHGEGEREE 66
DB 537 hvrgcktgcsaigentprkhtsrwshesrqse 568

RESULT 10
W62838 W62838 standard; Protein; 605 AA.
XX W62838;
XX
XX 27-OCT-1998 (first entry)
XX
XX Glycine max antimicrobial protein.
XX
XX antimicrobial protein; infestation; control.
XX
XX Glycine max.
XX
XX WO9827805-A1.
XX
```

```
PD 02-JUL-1998.
XX
XX 22-DEC-1997; 97MO-AU00874.
XX
XX 20-DEC-1996; 96AU-0004275.
XX
XX (REFR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
XX Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX
XX WPI; 1998-377279/32.
XX
XX Novel anti-microbial protein from e.g. Macadamia integrifolia -
XX useful for controlling microbial infestations of plants or mammals
XX
XX Claim 1; Page 63-65; 96pp; English.
XX
XX "The sequence is that of an antimicrobial protein which can
XX be used to control microbial infestations in plants and mammalian
XX animals.
XX
XX Sequence 605 AA;

Query Match 19.1%; Score 73; DB 19; Length 605;
Best Local Similarity 30.3%; Pred. No. 0.47;
Matches 20; Conservative 6; Mismatches 14; Indels 26; Gaps 3;

QY 15 QCVORCQERPRY-----SHARC-----VQEC-----RDDQOQHGRHEQ 48
DB 37 kclqscnsrdsyngqacharcmlkvekecegeglprprpqrphetepqpygeked 96
QY 49 EEOGR 54
DB 97 edeqpr 102

RESULT 11
Y40999 Y40999 standard; Protein; 605 AA.
XX
XX Y40999;
XX
XX 06-DEC-1999 (first entry)
XX
XX Soybean beta-conglycinin protein sequence.
XX
XX Peanut; allergen; Ara H 1; IgE; immunoglobulin E; epitope; Ara h 3;
XX allergic reaction; soybean; beta-conglycinin.
XX
XX Glycine max.
XX
XX WO9945961-A1.
XX
XX 16-SEP-1999.
XX
XX 12-MAR-1999; 99WO-US05494.
XX
XX 12-MAR-1998; 98US-0077763.
XX 11-MAR-1999; 99US-0077763.
XX
XX (UYAR-) UNIV ARKANSAS.
XX
XX Burke W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
XX Sampson H, Compadre CM, Huang SK, Maleki SJ, Koppen RA;
XX
XX WPI; 1999-551218/46.
XX
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
XX animal from allergic reaction -
XX
XX Disclosure; Fig 33A-B; 193pp; English.
XX
```



```

DT 10-APR-2000 (first entry)
XX
XX Human ovarian tumor EST fragment encoded protein 79.
DE
XX Expressed sequence tag; EST; human; ovarian tumor; anticancer;
KM gene therapy; treatment.
XX
XX Homo sapiens.
OS
XX DE19817557-A1.
PN
XX 21-OCT-1999.
PD
XX 09-APR-1998; 98DE-1017557.
PF
XX 09-APR-1998; 98DE-1017557.
PR
XX 09-APR-1998; 98DE-1017557.
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
PI
XX WPI; 1999-591920/51.
DR
XX N-PSDB; Z77484.
DR
XX
XX New nucleic acid sequences expressed in ovarian, and some other, cancer
PT tissues, and derived polypeptides, for treatment of ovarian cancer and
PT identification of therapeutic agents
XX
XX Claim 25; Page 276; 310pp; German.
PS
XX
XX This invention describes novel nucleic acid (cDNA) sequences (A) which
CC have anticancer activity and are highly expressed in ovarian tumor
CC tissue (and some also in testis and breast cancer tissue). The products
CC of the invention can be used for gene therapy. (A) are used (1) for
CC recombinant expression of polypeptides (B) and (1i) to isolate complete
CC genes. (B) are used (1) to identify agents suitable for treatment of
CC ovarian cancer; (1i) directly for treating this form of cancer (including
CC expression from gene therapy vectors) and (1ii) for generation of
CC specific antibodies. (A) are identified by assembling ESTs (expressed
CC sequence tags) from a particular tissue type before comparison of
CC expression patterns. This allows a significantly longer fragment of the
CC gene to be revealed, so should reduce the number of failures associated
CC with the fact that ESTs from different libraries may represent different
CC parts of the same unknown gene, distorting the estimated frequency of
CC occurrence in a particular tissue. Y76505-Y7638 represent protein
CC fragments encoded by the human ovarian tumor cDNA library derived EST
CC fragments represented in Z77450-Z77572.
XX
XX Sequence 156 AA;
SQ

```

Query Match 18.3%; Score 70; DB 20; Length 156;  
 Best Local Similarity 28.6%; Pred. No. 0.24;  
 Matches 18; Conservative 11; Mismatches 28; Indels 6; Gaps 2;

```

QY 7 RRGSHLQCCVQRCRQRRPRYSHA----RCVQECRDDQOQHGHDEDEGGRGWHGEG 62
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 70 rsgskartpqlflqqlqfigngcwrcwlaarehpgqgaqseegq--egq 127
   : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 ERE 65
   : : : : :
DB 128 qee 130

```

RESULT 15  
 Y32218  
 ID Y32218 standard; Protein; 1122 AA.  
 AC Y32218;  
 XX  
 XX 15-FEB-2000 (first entry)  
 DT  
 XX Drosophila PER B protein.  
 DE

```

XX
XX PER B, period; perl gene; transcription factor; circadian rhythm;
KM jet lag; sleep disorder; depression; seasonal affective disorder;
KM fertility; therapy.
XX
XX Drosophila melanogaster.
OS
XX W09957137-A1.
PN
XX 11-NOV-1999.
PD
XX 06-MAY-1999; 99MO-US10072.
PF
XX 07-MAY-1998; 98US-0084610.
PR
XX (HARD ) HARVARD COLLEGE.
PA
XX
XX Weitz CJ, Gekakis N, Staknis D;
PI
XX WPI; 2000-052938/04.
DR
XX
XX Novel heterodimeric composition for identifying modulators used in
PT diagnosing and treating circadian clock disruption disorders
XX
XX Disclosure; Fig 31; 96pp; English.
PS
XX
XX This sequence represents the Drosophila melanogaster PER protein
CC PER B. Splice variants are PER A (see Y32217) and PER C (see
CC Y32219). PER forms a heteromeric complex together with TIM (see
CC Y32221) protein and has a biological activity which inhibits
CC transcription of the perl gene when the CLOCK protein is present in
CC combination with BMAL1 protein (see Y32209). The invention is
CC based on the discovery of the transcriptional mechanism regulating
CC genes responsible for the establishment and/or maintenance of the
CC circadian clock, and provides an assay for novel drugs aimed at
CC restoration of a normal circadian cycle, the drugs being modulators
CC of BMAL1-CLOCK-mediated transcription of E-box-linked genes. The
CC drugs are used to treat conditions such as jet lag, sleep disorders,
CC depression (seasonal affective disorder) and infertility. The
CC invention also provides BMAL1 and CLOCK proteins with which to
CC stimulate the transcription of an E-box-linked gene which regulates
CC the circadian clock.
XX
XX Sequence 1122 AA;
SQ

```

Query Match 18.3%; Score 70; DB 21; Length 1122;  
 Best Local Similarity 36.2%; Pred. No. 2.2;  
 Matches 21; Conservative 9; Mismatches 14; Indels 14; Gaps 4;

```

QY 22 QERPRYSHARCVOE--CRDDQ-OQHGRH-----EQEEDGGRGWHGGEREE 66
   : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 103 eekprpsqfvcvqqlcrtelqdgqhgndhsepaqlqgseedsq--gsesadrlve 159
   : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: March 1, 2001, 16:09:35  
 Job time: 1327 sec